Geospatial Problem Solving

Attribute Tables and Geodatabases

Image from NASA:
https://commons.wikimedia.org/wiki/File:Earth_Western_Hemisphere_transparent_background.png#filelinks

WV VIEW

AmericaView
Empowering Earth Observation Education
AmericaView.org - Est. 2003

In this module, I will provide an in-depth discussion of attribute tables and geodatabases to expand upon our earlier discussion.

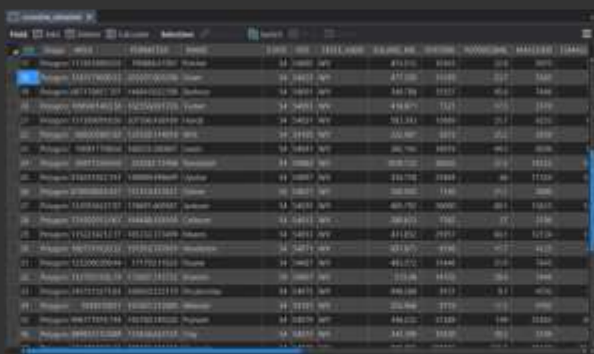
Module Topics

1. Attribute table structure
2. Field types
3. Adding fields
4. Geometric calculations
5. Table and spatial joins
6. Attribute table summarization and descriptive statistics
7. Aspatial queries
8. Spatial queries
9. File geodatabases, feature classes, and feature datasets
10. Domains
11. Subtypes
12. Relationship classes

These are the topics covered in this module.

Attribute Tables

Attribute Tables



Columns = fields or attributes; info stored for each record

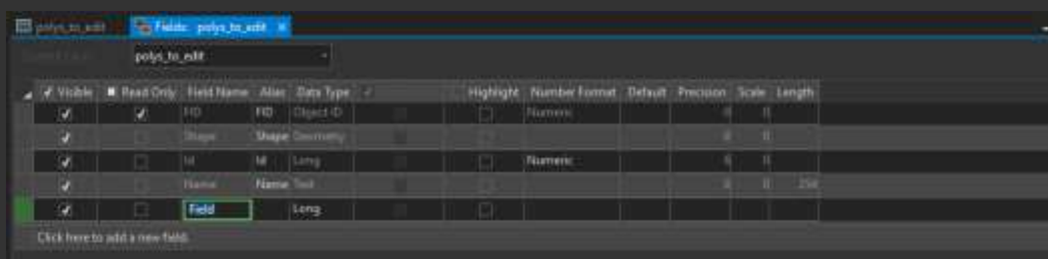
Rows = Records; one record for each vector feature

- All shapefiles and feature classes have an attribute table

Attribute or aspatial information is stored in the attribute table. In geospatial science, columns are commonly referred to as fields or attributes while rows are called records. Each row or record in the table stores all the aspatial data associated with one map feature, such as one county. Each column stores a specific piece of information for each feature, such as the county population.

Creating a new field

- Must define field name (less than 10 characters)
- Can provide an alias
- Must define data type

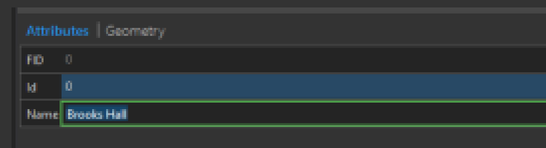


5

When you create a new field you can define the field name, which must be less than 10 characters for shapefiles and geodatabases. To make fields more interpretable, you can also include aliases, which provide more interpretable names or descriptions. You must also define the field type, which we will discuss next. You can also make fields not visible or read only if you don't want to allow editing. Note that the feature ID, or FID, will be defined as read only since you don't want to edit or change this unique identifier.

Adding Fields

❖Make sure to define the correct type



Data Type	Description
Short Integer	Integer values between -32,768 and 32,767
Long Integer	Integer values between -2,147,483,648 and 2,147,483,647
Float	Decimal values with 1-6 decimal places
Double	Decimal values with more the 6 decimal places
Text	Text strings
Date	Data info (mm/dd/yyyy)

6

When creating a new field, you will need to define the field type.

Short and long integer fields can only store integer or whole number values. Long integers can store a wider range of numbers than short integers.

If you need to maintain decimal values, you cannot use a short or long integer. Instead, you must use a float or double type. Float fields can store up to 6 decimal places while double can store more than 6.

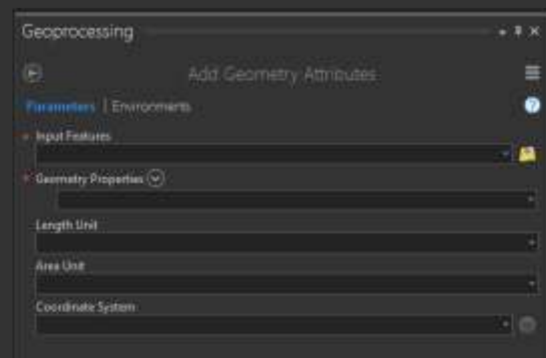
If you would like to store text or string data, you will do so in a text field. Note that numbers stored in text fields will be treated like strings. They will not be treated like numbers. For example, you will not be able to perform mathematical operations on these data.

Date fields store dates and will treat the information as dates.

Add Geometric Fields

❖ Use Add Geometry Attributes Tool

Points	Lines	Polygons
X-coordinate	Start-X	Area
Y-coordinate	Start-Y	Perimeter
Z-coordinate (if z-enabled)	End-X	Centroid-X
M-coordinate (if m-enabled)	End-Y	Centroid-Y
	Mid-X	
	Mid-Y	
	Length	
	Bearing	



7

Different geometric measures can be calculated and stored in the attribute table for points, lines, and polygon data.

For point data, it is possible to store the x and y coordinate. If the feature is Z-enabled, or contains elevation or height information, this value can also be calculated. If it is M-enable, the M coordinate can be calculated. M-coordinates generally relate to positions along a network.

For lines, you can calculate the x and y coordinates for the start, end, and middle of the line segment. You can also calculate the length and bearing. The bearing is the compass direction of the line and is defined as the direction of travel from the start to the end node.

For polygons, you can calculate the area and perimeter. The perimeter is the length of all the lines that make up the polygon. You can also calculate the centroid, or geometric center, coordinates. Note that centroid coordinates can sometimes be misleading. For example, for irregular or elongated shapes, the centroid may not actually occur within the polygon. For a "C" shaped feature, the centroid would occur somewhere in the middle of the "C" as opposed to within the area that makes it up. So, be careful when using centroid measures.

Joining non-spatial data and spatial data

Join

- Non-spatial data can be joined to spatial data as long as there is a common field between the data.
- Later, the joined fields can be used for analysis or to symbolize the features.

Relate

- New fields are not added to table. However, a relationship is established.

8

We discussed table joins in a prior lecture. Here I will provide a refresher.

Joins can be used to associate tables with spatial layers. If a table and a vector layer share a common field, then they can be joined based on this shared attribute. This is similar to the concepts of foreign and primary keys in relational databases.

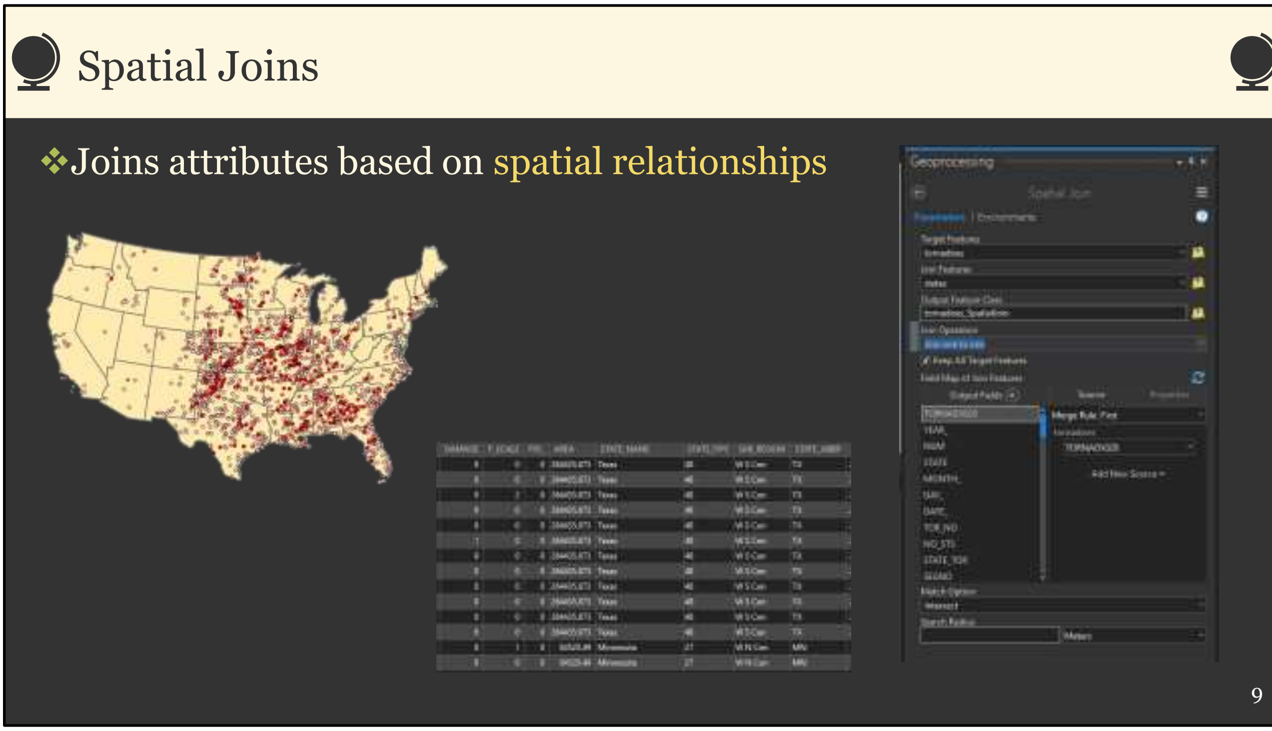
For example, if a layer of county boundaries contains the county name or FIPS code in its attribute table, this could be used to join a table to the features if the table also contains the name or FIPS code. Note that table joins are generally case sensitive and the shared fields must have the same data type.

Joins are just temporary. If you would like to permanently add the new columns to the attribute table, you will need to make a copy of the data.

A relate is similar to a join; however, the fields from the related table are not temporarily added to the feature attribute table. Instead, a relationship is established between the features, but the data are maintained as separate files.

I generally prefer to use join as opposed to relate.

Since both of these operations are just a temporary association, they can be removed. So, if you would like to remove the joined fields from the table, you can do so without any issues.



Since our vector data have both spatial and aspatial components, joins can also be performed using the spatial information as opposed to a common attribute field.

In the example in this slide, the points represent tornado locations and the polygons represent states. Using a spatial join, it would be possible to add the attributes associated with each state to the tornadoes occurring in the state using spatial co-occurrence as opposed to a shared attribute column. This is exactly what a spatial join is meant to accomplish.

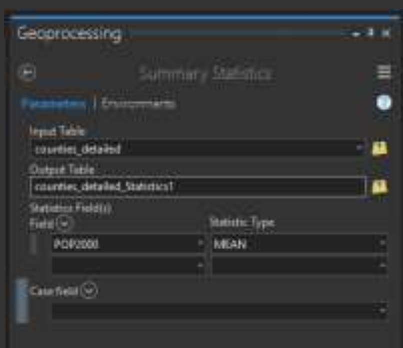
Table joins do not require spatial information. They could be performed using just tables without any associated map feature. In contrast, spatial joins rely on the spatial overlay of features.

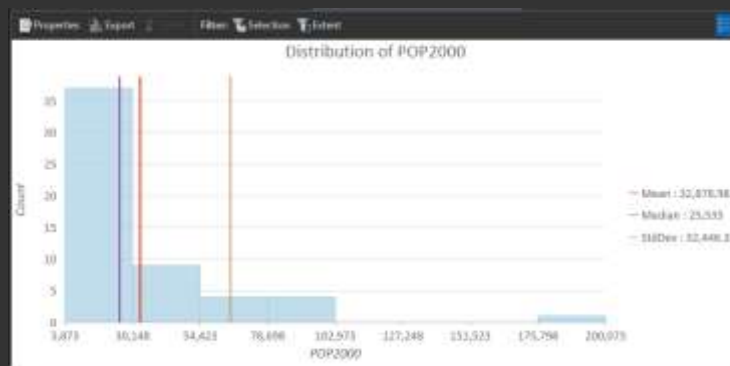
When a spatial join is performed, a new output is commonly generated as opposed to a temporary association, as is the case with a table join.

In this example, the output would be the tornado point locations with an attribute table containing the original attributes for the tornadoes and the attributes for the state in which the tornado occurred.

Field Summary Statistics

- Statistics for numeric fields





10

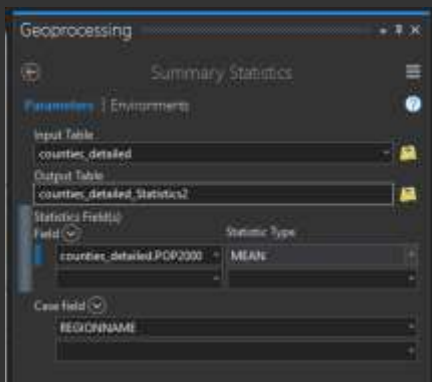
It is possible to obtain summary statistics as tables and charts for attribute fields.

You can also obtain summaries by case or group using a categorical attribute.

For example, you could summarize the mean percent clay content by soil type as opposed to a single summary for all soil types.

Field Summary Statistics

- Statistics for numeric fields



OBJECTID	WV_Deer_Harvest200-	FREQUENCY	MEAN_counties_detailed.POP2000
1	Eastern Panhandle	3	44346
2	Hatfield-McCoy Mo...	5	31301.8
3	Metro Valley	4	93600.75
4	Mid-Ohio Valley	7	23249.142857
5	Mountain Lakes	7	15542.285714
6	Mountaineer Country	7	39357
7	New River/Greenbrie...	8	38106.375
8	Northern Panhandle	6	28057.5
9	Potomac Highlands	8	15519.875

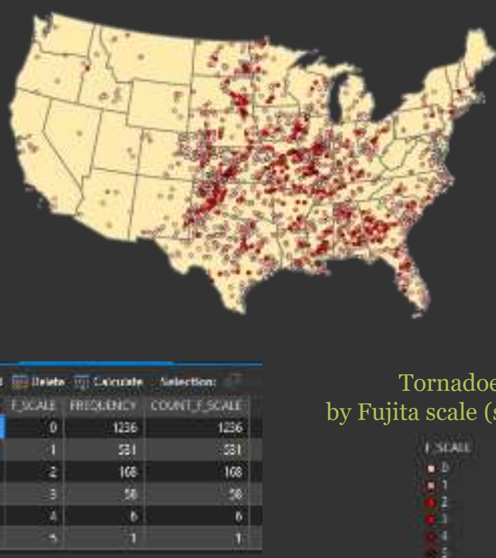
Click to add new row.

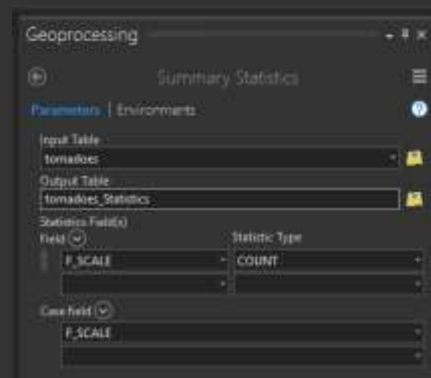
11

Here is an example of obtaining the average population of counties for different regions. Note that the mean statistic is being returned for a population field and that the case or group field is defined relative to a region name.

The result will be a table containing the region name, the number of counties in each region, and the returned statistic.

Summarize Fields





Field: Add Delete Calculate Selection:

OBJECTID	F_SCALE	FREQUENCY	COUNT_F_SCALE
1	0	1236	1236
2	1	531	531
3	2	168	168
4	3	58	58
5	4	6	6
6	5	1	1

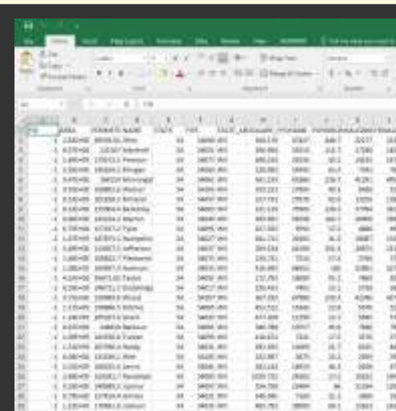
Tornadoes
by Fujita scale (severity)

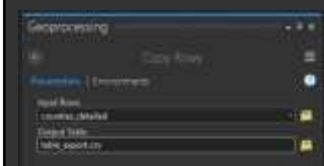
12

Here is another example where the number of tornadoes of each Fajita scale category have been counted.

Export an Attribute Data

- Types include text files, comma separated values (CSV), geodatabase tables, and dBASE tables
- Can analyze aspatial or tabulated data outside of GIS software





13

If you would like to work with the tabulated data and do not need the spatial component, attribute tables can be exported for use in data analysis or statistical software packages using a variety of formats, such as text files, comma separated values (or CSV) files, and database (or dBASE) files.

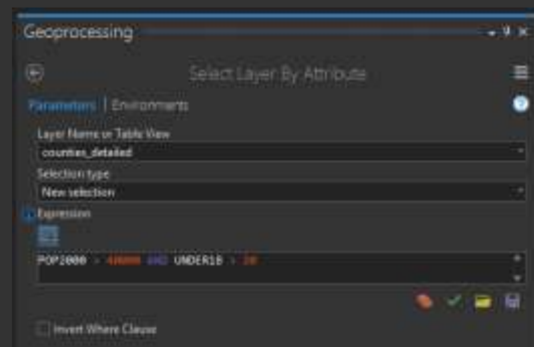
Table or Aspatial Queries

Logical Operators

Equal: =
Not Equal: <>
Greater Than: >
Greater Than or Equal To: >=
Less Than: <
Less Than or Equal To: <=

Boolean Connectors

AND (A and B)
OR (A or B, including A and B)
NOT (A but not B)
XOR (A or B, but not A and B)



Population >= 20,000 AND Region = Midwest

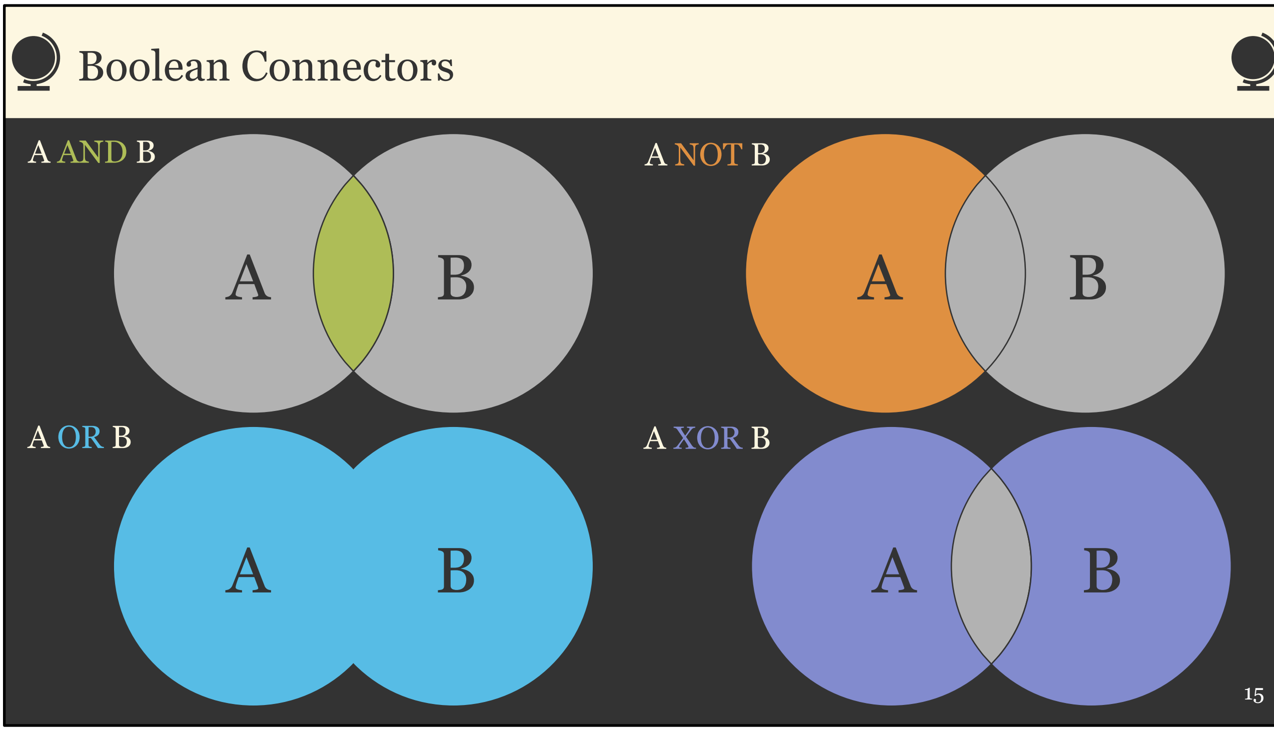
14

Data can be queried based on the information stored in the attribute table. Such queries are known as table or aspatial queries. These queries can be performed using just attribute information. The geometric information is not required to complete the query.

Queries make use of logical operators, such as equal to, not equal to, less than, and greater than or equal to. They can also make use of Boolean connectors if multiple criteria will be included.

In the example in this slide, counties are being queried. Specifically, the goal is to extract all counties that have a population greater than or equal to 20,000 and are in the Midwest region. Because AND is used here, both criteria must be met for a feature to be selected.

Once a query is executed, selected features can be written to a new file.



AND = Only areas that occur within A and B are selected (green).

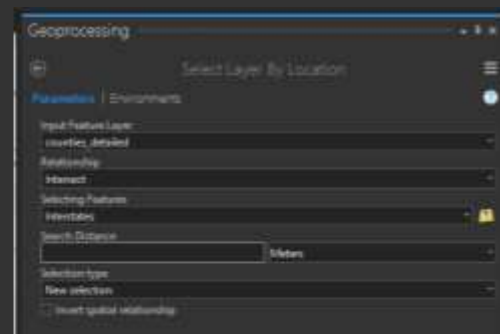
OR = All areas that are within A, B, or both and A and B are selected (blue).

NO = Areas that are in A but not B are selected (orange).

XOR is different from OR. Here, only areas that are in A or B, but not both A and B are selected (purple). So, OR will yield A or B, including A and B. In contrast, XOR will yield A or B, excluding A and B.

Select by Location

- Select based on spatial relationships to other geographic features



16

It is also possible to query features based on spatial relationships to other map features. Such queries are not possible for tabulated data since the geometric information is necessary to perform the analysis. So, this is unique to spatial data and highlights the value of having spatial information associated with tabulated data.

Using select by location, features can be selected based on their spatial relationship to other features.

The next two slides describe some of the available spatial relationships.

Select by Location

- Intersect = select input features that intersect the selecting features. This is the default.
- Within a Distance = select input features that are within the specified Euclidean distance of the selecting features.
- Contains = select any input feature that contains, even partially, a selecting feature.
- Completely Contains = select any input feature that completely contains a selecting feature.
- Within = select any input feature that is within, even partially, a selecting feature.
- Completely Within = select any input feature that is completely within a selecting feature.

17

Take some time to read through the spatial relationships presented here.

As an example, the intersect relationship can be used if you want to select features that intersect another set of features. For example, you could select all counties that intersect a particular highway. The county will be selected if a highway intersects it or passes through it.

Within can be used to select features that are within other features. For example, you could select all of the homes that occur within a certain voting district. In contrast, contains could be used to select all of the voting districts that contain certain homes.

Select by Location

Are identical To = select any input feature that has identical geometry to a selecting feature.

Boundary Touch = select any input feature that has a boundary that touches a selecting feature.

Share a Line Segment With = select any input feature that shares a line segment with a selecting feature.

Crossed by Outline Of = select any input feature that is crossed by the outline of a selecting feature.

Have Their Center In = select any input feature that has its center in the selecting features. The center for lines is defined as the midpoint and the center for polygons is defined as the centroid.

Here are some additional spatial relationships.

As an example, are identical to is used to find features in one layer that are identical to features in another layer. For example, if you had a polygon vector dataset of buildings representing the city in 1990 and another layer representing buildings in the city in 2010, you could use this tool to find the identical buildings, or buildings that haven't changed. You could then reverse the selection to find all new buildings.

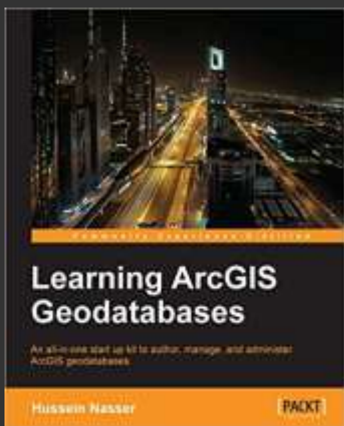
Take some time to read through these examples.

Geodatabases

gid integer	postgis_fi bigint	lat_y_dd numeric	long_x_dd numeric	status character varying (254)	nest_id bigint	geom geometry
1	1	[illegible]	[illegible]	ACTIVE NEST	1	[illegible]
2	35	[illegible]	[illegible]	INACTIVE LOCATION	35	[illegible]
3	36	[illegible]	[illegible]	ACTIVE NEST	36	[illegible]
4	37	[illegible]	[illegible]	ACTIVE NEST	37	[illegible]
5	40	[illegible]	[illegible]	ACTIVE NEST	40	[illegible]
6	47	[illegible]	[illegible]	ACTIVE NEST	47	[illegible]

19

Book Recommendation



https://www.packtpub.com/application-development/learning-arcgis-geodatabases

20

If you are interested in learning more about geodatabases in ArcGIS, I would recommend this text.

Why use a Geodatabase?

1. Organize and manage data
2. Make use of database management systems
3. Make use of structured query language (SQL)
4. On-the-fly topology
5. Topology between features
6. Use domains and subtypes
7. Migrate to an enterprise database

21

Other than storing your data in flat files or separate files, the information can be stored in a database structure.

Databases offer many benefits, including the ability to organize and manage your data, make use of database management systems, use Structured Query Language (SQL) for data manipulation and querying, model topological relationships on-the-fly and between layers, define domains and subtypes, and migrate data to an enterprise database.

The Relational Model

- Collection of interrelated tables in digital format
- Tables are connected via keys
- Primary Key = one or more attributes whose values can uniquely identify a record in a table
- Foreign Key = one or more attributes that refer to a primary key in another table
- Each table can be prepared, maintained, and edited separately
- Tables can be related

22

Another benefit of geodatabases is the ability to relate data layers and tables. For example, you can make associations using primary and foreign keys. If a database contains vector features representing counties and tables of information for each county, the county name could be used as a foreign key to join the table to the vector features using the county name in the attribute table associated with each map feature. Even though the layers are now associated, each data layer or table can be prepared, maintained, and edited separately.

This is the relational model in action. Databases allow for multiples datasets to be stored separately but associated within the database.

The Geodatabase

File Geodatabases

- ❖Edited by a single user and do not support versioning
- ❖Usually hosted and used on a local machine

Enterprise Geodatabase

- ❖Edited by multiple users and supports versioning
- ❖Host locally or in an enterprise database system (Oracle, Microsoft SQL Server, IBM DB2, Informix, PostgreSQL)

23

File geodatabases are designed to receive edits from only one user at a time. However, multiple users can access the data at once. These databases are generally hosted or stored on a local machine as opposed to on a server using an enterprise database management system.

Enterprise geodatabases are meant to be stored on a server using a database management system and allow for multiple editors and users at once.

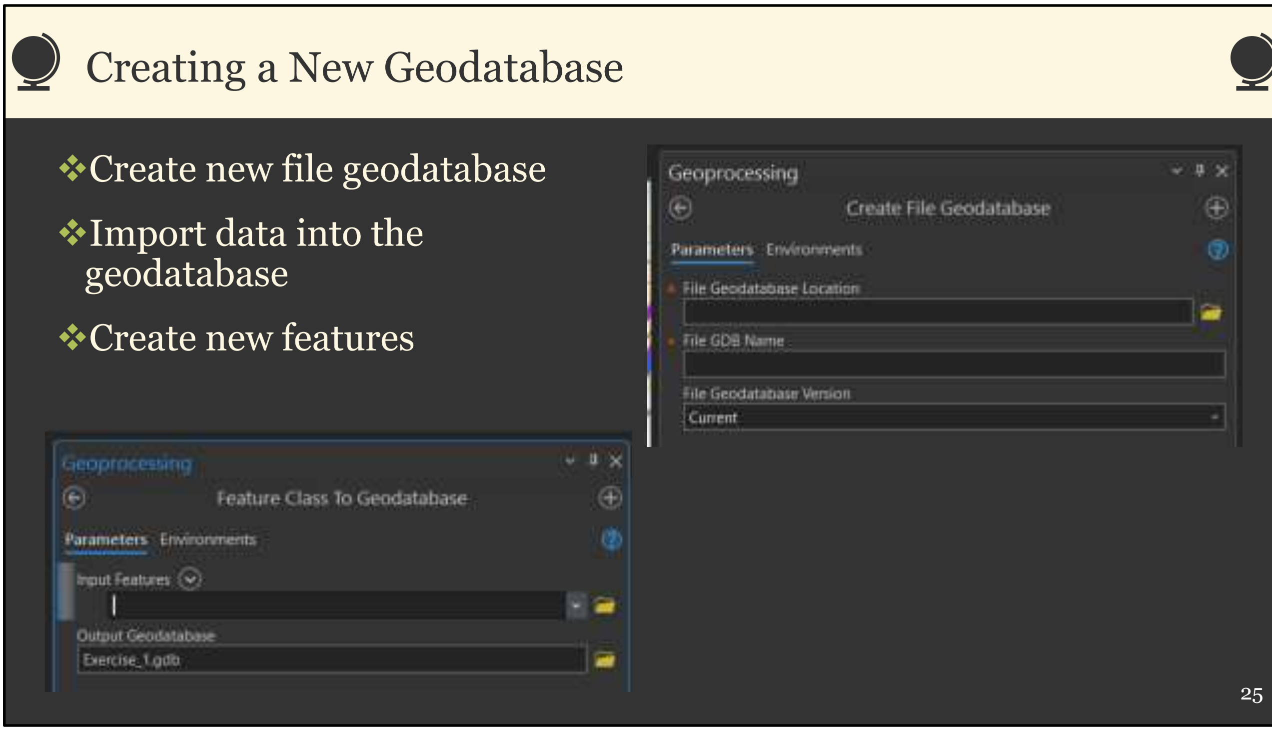
ESRI Geodatabases

	File GDB	Enterprise GDB
Storage Format	Folder of binary files	Database management system (Oracle, Microsoft SQL Server, IBM DB2, Informix)
Storage Capacity	1 TB per table	Depends
Supported OS	Any platform	Depends
Number of users	Single editor, multiple users	Multiple editors, multiple readers

https://pro.arcgis.com/en/pro-app/help/data/geodatabases/overview/types-of-geodatabases.htm

24

This table provides a comparison of the two types of databases.



A new file geodatabase can be created in ArcGIS Pro using the Create File Geodatabase tool.

Once a database is created, you can create new empty feature classes within it or import data layers, such as feature classes, shapefiles, tables, and raster grids. Data can be imported with the Feature Class to Geodatabase tool.

You can also create file geodatabases within the Catalog window or pane.

Geodatabase Components

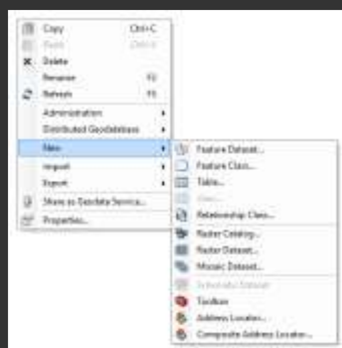
- Feature classes
- Feature datasets
- Mosaic datasets
- Raster catalog
- Raster dataset
- Schematic dataset
- Tables
- Toolboxes
- Attachments
- Feature-lines annotations
- Geometric networks
- Network datasets
- Parcel fabrics
- Relationship classes
- Terrains
- Topologies

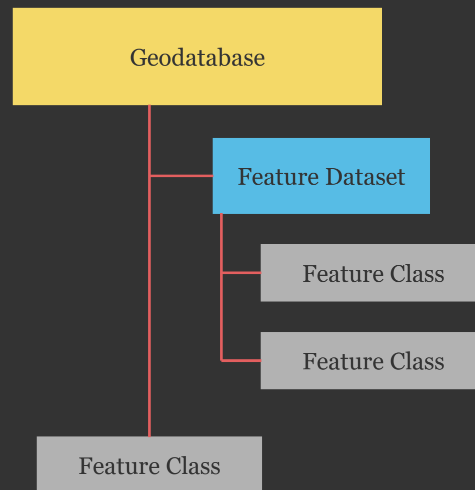
26

A variety of different features can be stored in a geodatabase. Some of these features will be described in subsequent slides.

Feature Class

- Collection of features that share the same geometry (points, lines, or polygons) and spatial reference





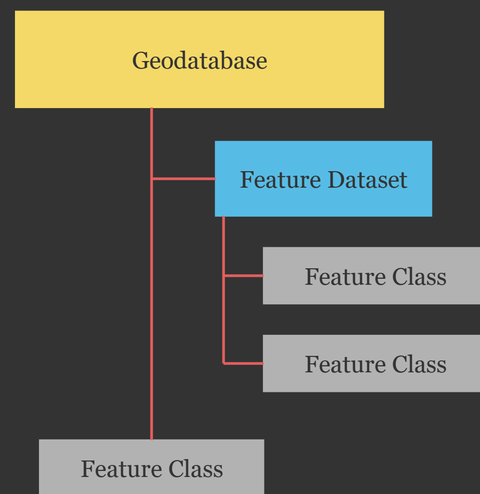
27

Inside a geodatabase, vector data are stored as feature classes. All features inside of a feature class must share the same geometry (points, lines, or polygons). A spatial reference must be applied to a feature class.

This is an object-based model since the geometry and attribute data are stored in the same file. In a feature class, the geometry is stored as a column in a table as a binary large object or BLOB.

Feature Datasets

- Collection of feature classes that have the same spatial reference
- Can be used to organize the geodatabase
- Can control data access and editing privileges
- Can organize data for sharing
- May be necessary
  - Build topology between features
  - Create a network dataset
  - Create a terrain dataset
- Must define spatial reference upfront



28

Feature datasets are collections of feature classes that have the same spatial reference. You can think of feature datasets as subfolders within the database. They can be used to organize your data. For example, if you have a geodatabase of utility data for a city, you could organize the electrical, water, and communication data into separate feature datasets.

You can also use feature datasets to control data access and editing privileges. For example, you could allow all utility workers access to data, but only those working on the electrical utilities could edit data in the feature dataset containing the electrical utility information.

Sometimes feature datasets are required. Examples include when you want to build or define topological relationships between layers, create network datasets, or create terrain datasets.

All data in the feature dataset must have the same spatial reference.

Domains

- A property of a geodatabase shared by the individual datasets
- Created at the geodatabase level
- Can be used to specify allowed values or entries to a field
- Coded Values = discrete values that represent allowed entries
- Ranged Values = stipulates that values can only be entered within a given range

Code	Description
1	Poor
2	Fair
3	Average
4	Good
5	Excellent

29

Domains are defined at the database level and can be applied to multiple features stored in the database. They are used to define allowed values or entries to a field.

For example, if a field is meant to store a rating, you may only want the user to be able to enter between 0 and 5 stars.

Domains can also be used to define allowed categories. For example, a field of movie genres should only be able to store genres from a predefined list.

Coded values are generally used to set up allowed discrete entries or categories. For example, if you want to create a field that will store rock types for a geologic field collection, then you can predefine the different types that are allowed.

Ranged values are generally used when you want the user to be able to enter a number, but only within a defined range. For example, if a field is meant to store a percentage, then you would want the user to only be able to enter values between 0 and 100.

In the provide examples, coded values are being used to define the level of difficulty of a trail whereas ranged values are being used to specify the range of values allowed in a rating column.

Creating domains can be very useful. This will force the user to only enter valid data into a field.

Subtypes

- ❖Subsets of features in a feature class or table that share the same attribute
- ❖Used to categorize data
- ❖Use coded values

30

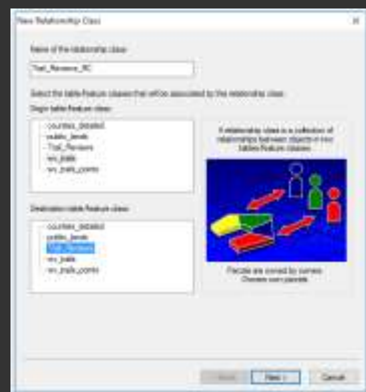
Subtypes represent features in a feature class or table that share the same attribute and are used to categorize data using coded values.

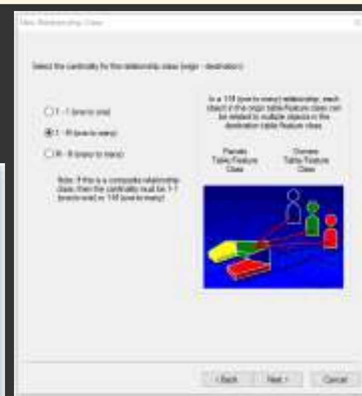
For example, in a layer that contains land parcels, subtypes could be defined relative to how the parcel is zoned (for example, public, commercial, or residential).

In the provided example, trails are being grouped into three subtypes as trail, motorized, and water.

Relationship Class

- Establishes a relationship between features in a geodatabase
- Define relationship using primary and foreign keys





31

Again, one of the benefits of using a geodatabase is the ability to relate tables and data using primary and foreign keys. In a file geodatabase, this is accomplished using a relationship class.

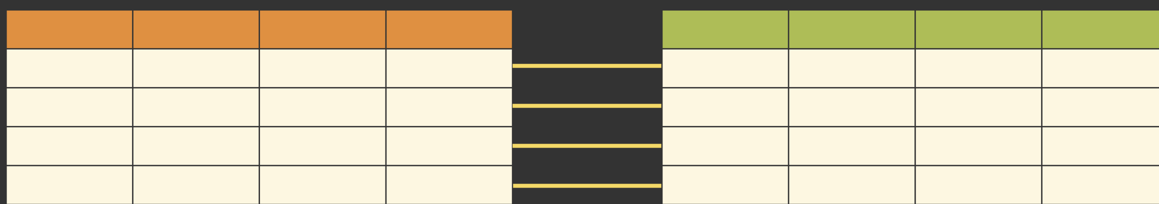
Types of Relationships

- One-to-One = one record in Table 1 is related to 1 record in Table 2
- One-to-Many = one record in Table 1 is related to many records in Table 2
- Many-to-One = multiple record in Table 1 is related to the same record in Table 2
- Many-to-Many = multiple records in Table 1 is related to multiple records in Table 2

32

Different types of relationships are available. Note that many of these types are not available when using joins and relates outside of a database, such as joining a table to a shapefile.

One-to-One

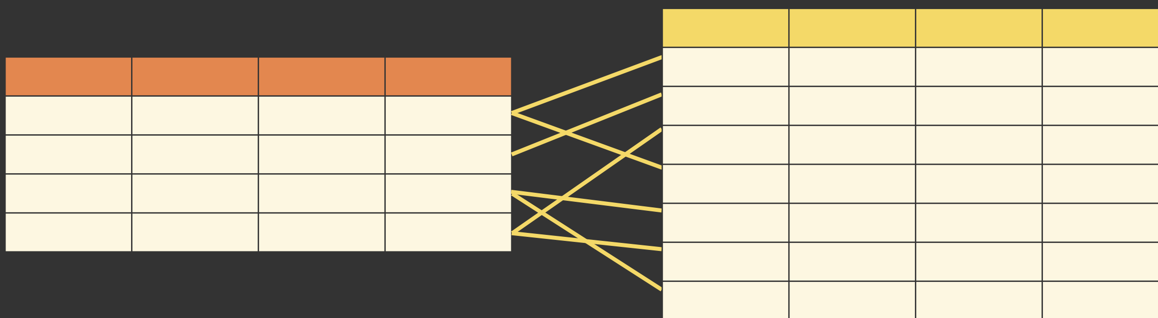


33

One-to-one relationships only allow for one record to be associated with one other record. This is what is commonly done when performing table joins and relates outside of a database.

As an example, one county could be associated with one record associated with that county from a table.

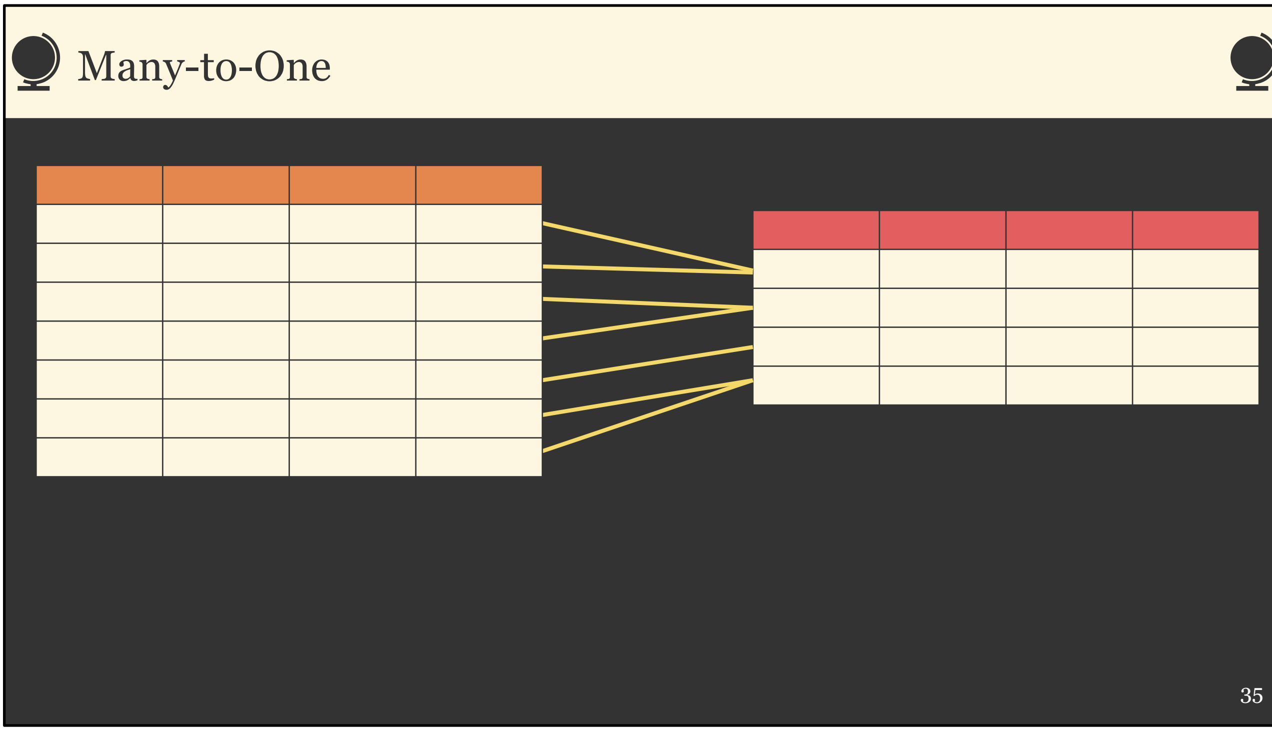
One-to-Many



34

It is also possible to associate one record to multiple records using a one-to-many relationship.

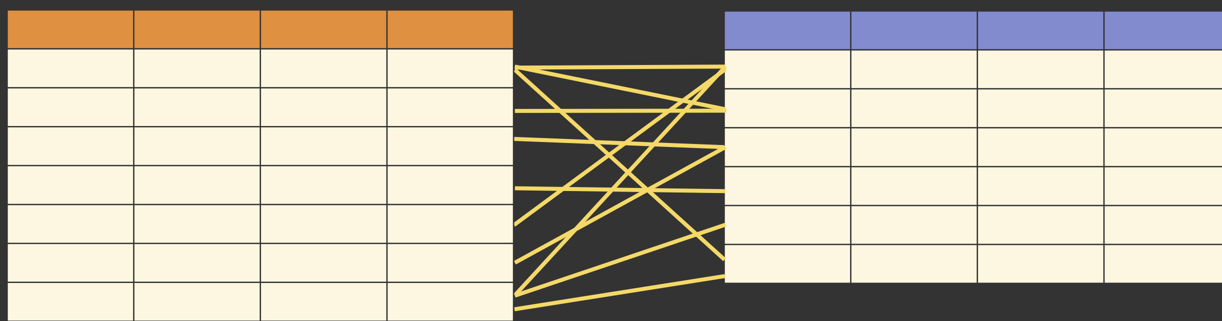
For example, one site could be associated with multiple field surveys from different dates.



Many records can be associated with a single record in another table using a many-to-one relationship.

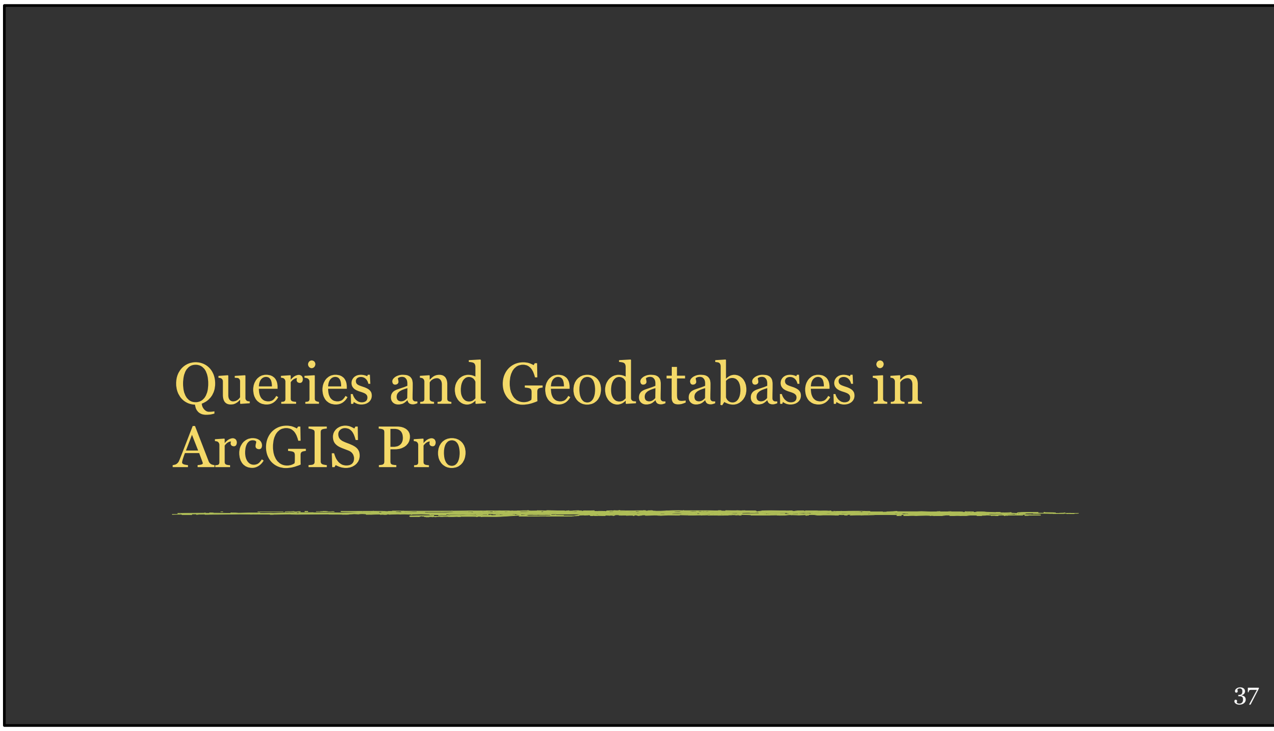
For example, each student in a table could be associated with a table of general information for students in the same academic year. So, freshman, sophomore, junior, senior, and graduate student information could be associated with each student of that academic class standing.

Many-to-Many



36

Many-to-many relationships allow for multiple records in one table to be associated with multiple records in another table.



The following exercises in ArcGIS Pro explore selection and working with geodatabases.

Attribute Information

1. Make sure the wv map is active.
2. Right-click on the summits layer in the Contents Pane then select Attribute Table to open the attribute table.
3. Use the attribute table to answer the following questions. Note that the number of records in the table is listed at the bottom of the window. If you hover over a column name, info about the column is displayed.

How may records are in the table?

What is the data type of the “ELEV_IN_FT” field?

Table Queries 1

1. Make sure the wv map and Map Tab are active.
2. Select the Select By Attribute icon in the Selection area of the Map Tab. 
3. Set the **Input Rows** parameter to summits.
4. Set the **Selection Type** to *New Selection*.
5. Enter the following query: Where [ELEV_IN_FT] is greater than [3500].

How many summits have an elevation greater than 3,500 feet?

Table Queries 2

1. Use table query methods to answer the following questions.

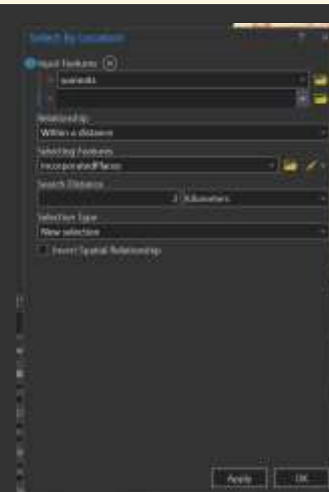
How many summits are in Randolph county?

Of all the summits in Randolph County, how many have an elevation greater than 3,800 feet?

40

Select by Location 1

1. Make sure the wv map and Map Tab are active.
2. Select the Select By Location icon in the Selection area of the Map Tab.
3. Set the **Input Features** parameter to summits.
4. Set the **Relationship** to *Within a distance*.
5. Set the Selecting Features to incorpratedPlaces.
6. Set the **Search Distance** to *2 kilometers*.
7. Set the **Selection Type** to *New selection*.
8. Do not select **Invert Spatial Relationship**.



How many summits are within 2 kilometers of an incorporated area?

The count of selected features for a layer is provided at the bottom of its attribute table. Or, you can change the Contents Pane to List by Selected. The number of selected features listed at the bottom of the map includes selections from all layers. !

41

Select by Location 2

1. Use Select By Location to answer the following questions.

How many summits are within 0.5 kilometers of a trail?

Do any of the higher education facilities not occur within an incorporated area?

Field Summary Statistics

1. Make sure the wv map is active.
2. Right-click on the summit layer and click Attribute Table to open its attribute table.
3. In the attribute table, right-click on the “ELEV_IN_FT” field.
4. Click on Statistics. Use the provided statistical information to answer the following questions.

What is the mean elevation of the summits?

What is the lowest elevation of the summits?

Field Summary Statistics by Group

1. Make sure the wv map is active.
2. Right-click on the summit layer and click Attribute Table to open its attribute table.
3. In the attribute table, right-click on the "ELEV_IN_FT" field.
4. Click on Summarize. This will open the Summary Statistics Tool.
5. Set the **Input Table** to summits and the **Output Table** to statsByCounty. Set the "ELEV_IN_FT" column as a **Statistics Field** and calculate the *Mean* **Statistic Type**. Set the **Case Field** to "COUNTY_NAM". Click Run to execute the tool.
6. Open the resulting table. The table will be listed in the Contents Pane under Standalone Tables. You can right-click on the table to open it. Use the table to answer the following question.

Which county has the highest mean elevation for its peaks (note some counties neighboring West Virginia are included in the dataset)?

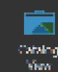
You can sort a table ascending or descending relative to a column by right-clicking on the column and selecting Sort Ascending or Sort Descending.

!

44

Make a File Geodatabase

1. Navigate to the Catalog View. If the Catalog View is not open, navigate to the View Tab and click the Catalog View button to activate it.
2. Click on the Databases option.
3. Within Databases, right-click and select New File Geodatabase
4. Create a new database called testDB.gdb in the day1 folder. Click Save to create the database.



The Catalog is available as both a View and a Pa .

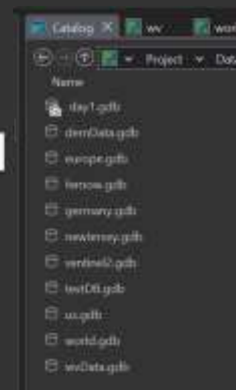
!

Database will be configured to store geologic data and field point measurements of rock type, strike, and dip.

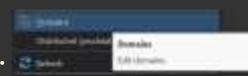
45

Create Coded Values Domain

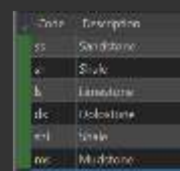
1. Double-click testDB.gdb in the Catalog View to open it. It should be empty.



2. Right-click in the empty space within the geodatabase and select Domains.



3. Create a Domain with the following characteristics.
   - **Domain Name** = *rockType*
   - **Description** = *Type of rock*
   - **Field Type** = *Text*
   - **Domain Type** = *Coded Values Domain*
   - **Split Policy** = *Default*
   - **Merge Policy** = *Default*
   - Fill out the Code/Description Table as follows:



   - Click Save in the Domains Tab to save the domain.

Create Range Domains

1. Using a similar process as outlined in the last slide, create a Range Domain called *dipVal* with a **Minimum** value of *0*, a **Maximum** value of *90*, and a **Field Type** of *Double*.
2. Using a similar process as outlined in the last slide, create a Range Domain called *strikeVal* with a **Minimum** value of *0*, a **Maximum** value of *360*, and a **Field Type** of *Double*.
3. Remember to Save your changes. The resulting domains should be configured as follows:

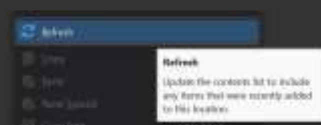
Domain Name	Description	Field Type	Domain Type	Split Policy	Merge Policy
dipVal	Rock dip	Double	Range Domain	Default	Default
rockType	Type of rock	Text	Coded Value Domain	Default	Default
strikeVal	Rock strike	Double	Range Domain	Default	Default

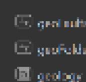
47

Import Into File Geodatabase

1. Navigate back to Databases area in the Catalog View.
2. Double-click testDB.gdb in the Catalog View to open it. It should be empty.
3. Right-click inside the database and select Import → Feature Class(es). This will open the Feature Class to Geodatabase Tool.
4. Use the folder icon next to **Input Features** to navigate to wvData.gdb. Select the geoFaults, geoFolds, and geology layers (you can hold down CTRL to select multiple layers). Click OK.
5. Click Run to execute the tool. 

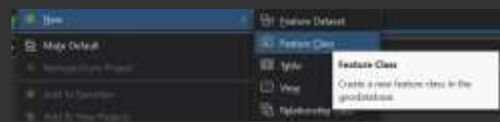
The three files should be added to the database. You may need to right-click and select refresh to see the changes.





Create a New Feature Class

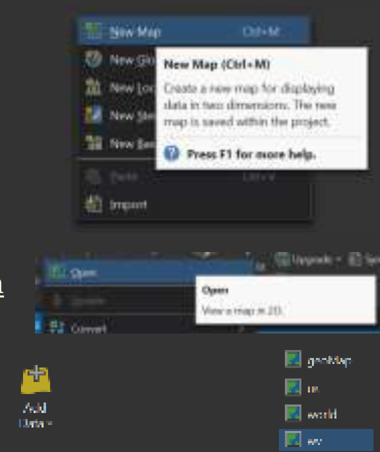
1. Navigate back to Databases area in the Catalog View.
2. Double-click testDB.gdb in the Catalog View to open it. It should be empty.
3. Right-click inside the database and select New → Feature Class.
4. This will Open the New Feature Class Pane. Fill out the pane as follows (there are multiple windows; click <u>Next</u> between windows and <u>Finish</u> in the final window; any options not listed can be left as the default):
   - **Name** = *fieldObservations*
   - **Alias** = *Rock Field Observations*
   - **Feature Class Type/Type of features stored in the feature class** = *Point*
   - Do not fill out the **Fields** options (you will do this later).
   - **Projection** = *Projected Coordinate Systems → UTM → North America → NAD 1983 (2011) UTM Zone 17N.*



49

Create a New Map

1. Navigate to the Maps area in the Catalog View. Three maps should already be available: us, world, and wv.
2. Right-click within the Maps area and select New Map.
3. The new map will automatically load. You can close it and navigate back to the Catalog View.
4. Right-click the map and select Rename. Rename the map *geoMap*.
5. Either double-click or right-click on geoMap and select Open to open a new map view.
6. Add all of the layers from the testDB.gdb using the Add Data button (you can hold down CTRL to select multiple layers at once).



Add Fields to Feature Class

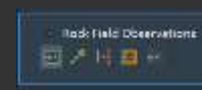
1. With the geoMap map active, right-click on the Rock Field Observations layer in the Contents Pane and select Attribute Table.
2. In the attribute table, click on Add . This will open the Fields view for the table.
3. Add the following fields with the following characteristics. Make sure to Click Save in the Fields Tab to save the changes.

- "Rock Type"
  - **Field Name** = *rType*
  - **Alias** = *Rock Type*
  - **Data Type** = *Text*
  - **Domain** = *rockType*
- "Strike"
  - **Field Name** = *strike*
  - **Alias** = *Strike*
  - **Data Type** = *Double*
  - **Domain** = *strikeVal*
- "Dip"
  - **Field Name** = *dip*
  - **Alias** = *Dip*
  - **Data Type** = *Double*
  - **Domain** = *dipVal*

51

Test Feature Class

1. Navigate back to the geoMap map.
2. In the Edit Tab, select Create . This will open the Create Feature Pane.
3. In the Create Feature Pane, click on the Rock Field Observations layer and the point option.
4. Add a point to the map (doesn't matter where) by clicking in the map space.
5. In the Edit Tab, click on Attributes. This will open the Attributes Pane. Make sure you can fill out the "Rock Type", "Strike", and "Dip" fields and that the domains are honored. Click Apply to save the changes.
6. To stop editing, select Save in the Edit Tab.

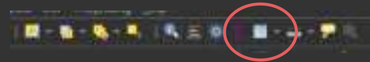


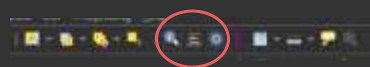
52

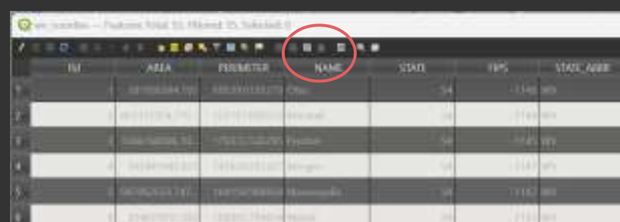
Queries in QGIS

53

Attribute Table and Field Calculator

- Open Attribute Table 
- Right-Click → Open Attribute Table

- Open Field Calculator 

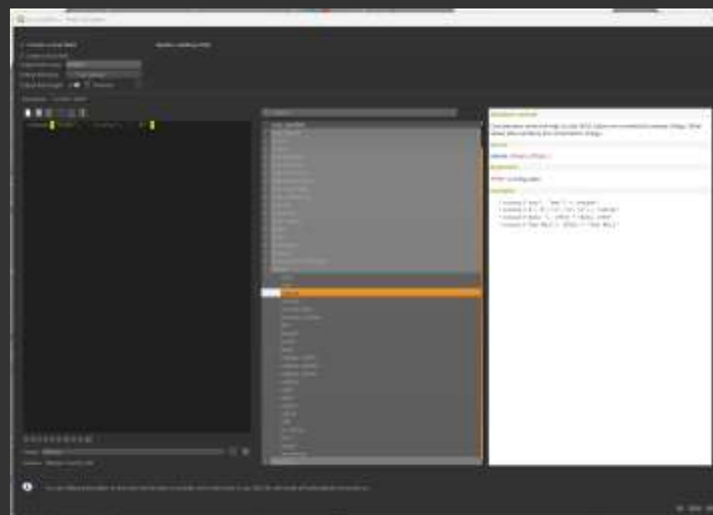


54

This slide describes different ways to open an attribute table in QGIS.

Create New Field

- Field Calculator
- Name field
- New or update existing
- Define field type
- Define expression
- Perform calculation



55

When performing a field calculation, you must name or specify the field, use a new field or update an existing field, define the field type, and build an expression to populate the field. If the expression is valid, you can then perform the calculation.

Building an Expression

- ID, feature, geometry, row_number
- Fields and values
- Geometric measures
- Math operations
- String manipulation
- Math, logical, and comparison operators
- File and file paths
- Etc.

56

A variety of different operators, data elements, and functions are available to perform field calculations.

Math, Logic, Comparison

- Addition: +
- Subtraction: -
- Multiplication: *
- Division: /
- Modulus: %
- Exponentiation: ^
- Less than: <
- Greater than: >
- Less than or equal to: <=
- Greater than or equal to: >=
- Not equal to: <>
- Equal to =
- AND
- OR
- NOT
- BETWEEN
- NOT BETWEEN
- IN
- IS

57

This slide describes the available math, logic, and comparison operators available when performing field calculations.

Aggregation

- count()
- count_distinct()
- count_missing()
- mean()
- median()
- minimum()
- maximum()
- iqr()
- q1()
- q3()
- range()
- stdev()
- sum()
- majority()
- minority()

58

This slide lists functions available for aggregating data.

Strings

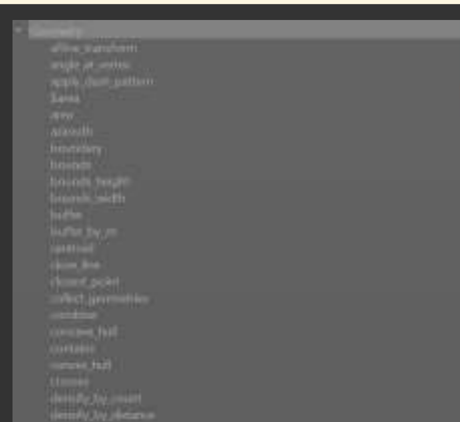
- Fields = “”
- Constant strings = ‘ ’
- concat()
- format_date()
- format_number()
- length()
- lower()
- upper()
- title()
- left()
- right()
- substr()
- trim()
- replace()
- Etc.

59

These slide lists functions available for manipulating strings.

Geometry

- $area, area()
- $perimeter, perimeter()
- $length, length()
- $x, x()
- $y, y()
- centroid()
- x_max()
- x_min()
- y_max()
- y_min()
- bounds()
- concave_hull()
- convex_hull()
- minimal_circle()
- start_point()
- end_point()
- Etc.



This slide lists functions available for working with and calculating geometric attributes.

Files and Paths

- base_file_name() = file name without full path or suffix
- file_exists() = tests whether file exists
- file_name() = file name with extension
- file_path() = path without file name and suffix
- file_size() = returns file size in bytes
- file_suffix() = returns just suffix or extension

61

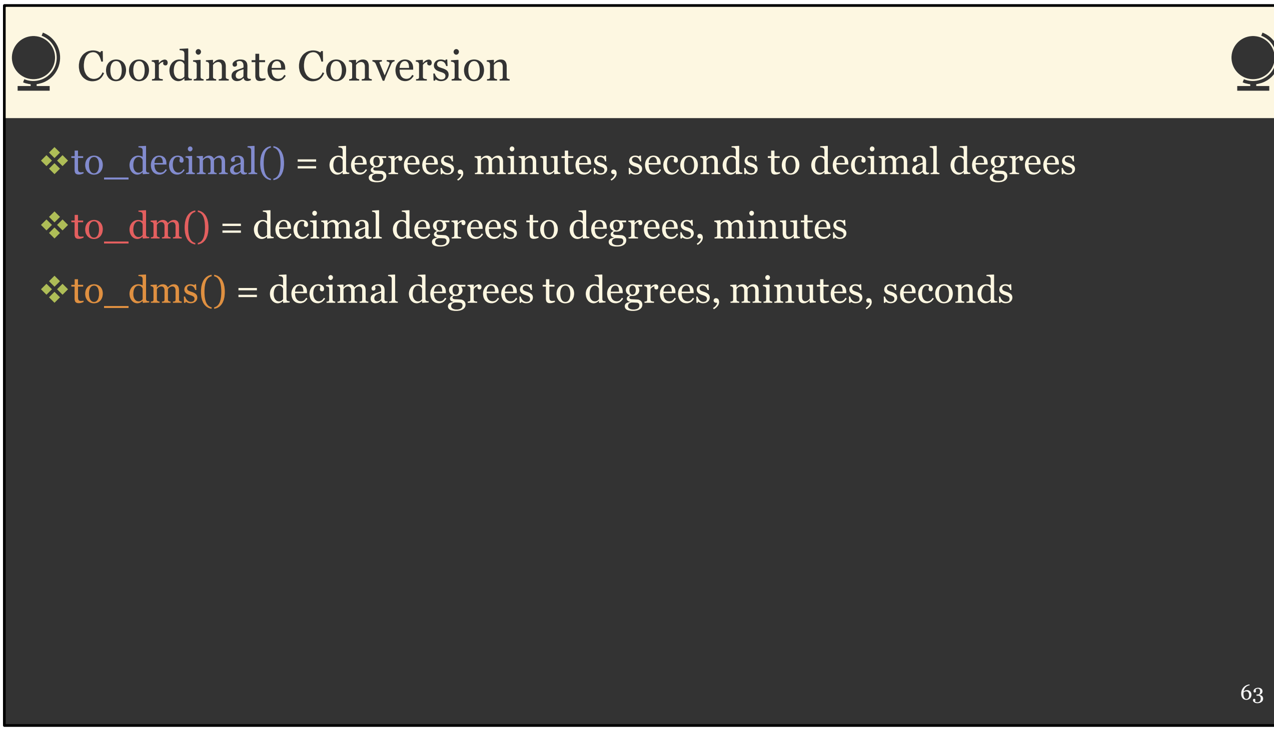
This slide describes functions for working with file names and paths.

Numeric and String Conversion

- to_string() = convert number to string
- to_int() = convert string to integer
- to_real() = convert string to real number
- to_time() = convert string to time
- to_date() = convert string to date
- to_datetime() = convert string to date and time

62

This slide describes functions for converting between different data types.



This slide describes functions for converting coordinate measurements.

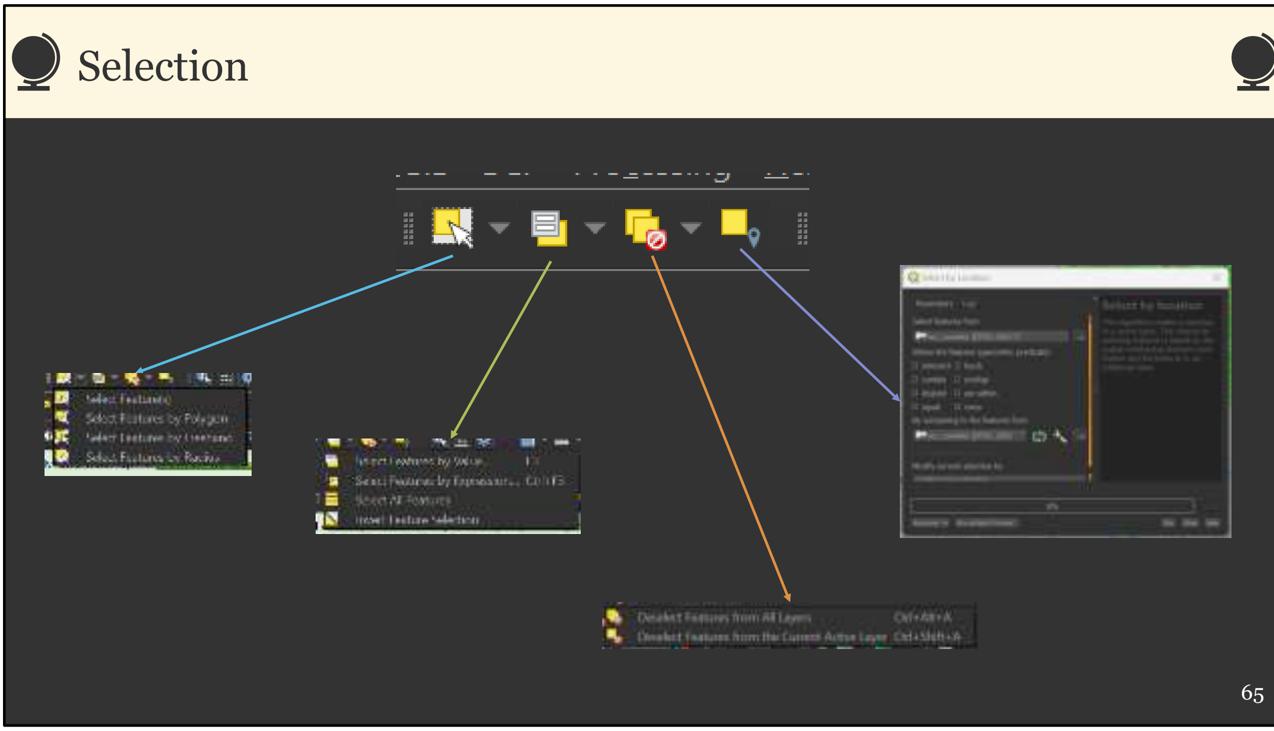
QGIS Attribute Table Calculations

- Use the utah project.
- Perform the following attribute table calculations.

1. Add the longitude and latitude of the towns points to the files attribute table in decimal degrees format.
2. Add the UTM Zone 12 North easting and northing coordinates of the town points to the files attribute table in meters units.
3. Calculate the length of trails in the trails layer in miles units.
4. Create a copy of the NAME field in the counties layer that is in title case.

64

This exercise asks you to perform table calculations in QGIS.



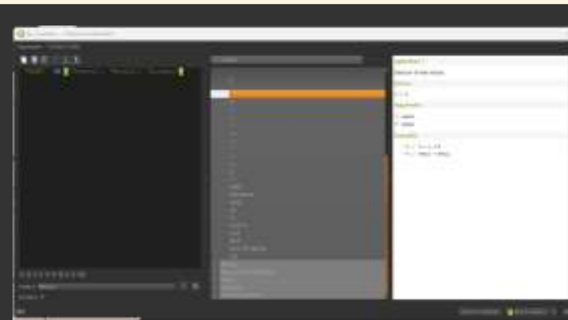
Different selection methods are available in QGIS and are introduced on the following slides.

Select Features by Values

66

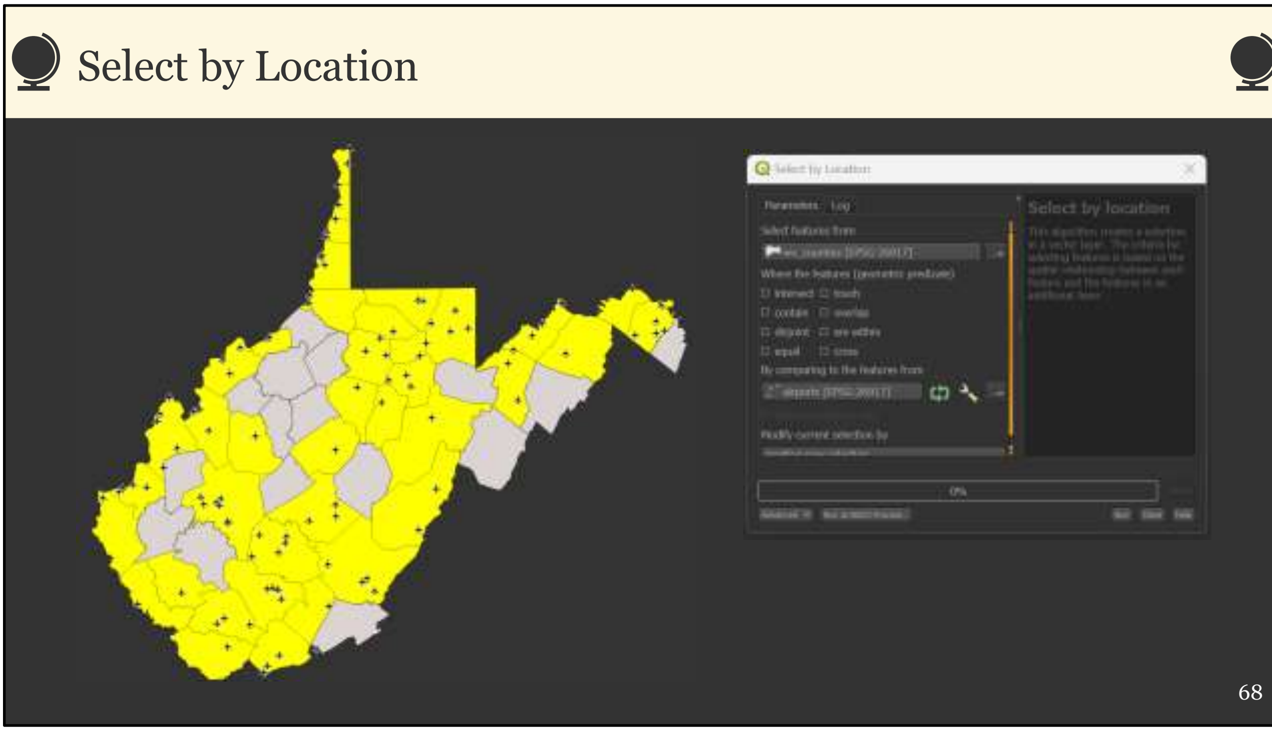
Select Features by Values allows you to perform attribute table queries by building expressions using a wizard interface.

Select Features by Expression



67

Select Features by Expression allows you to perform queries using an expression builder and available values, attributes, and functions.



Select by Location allows for defining location-based queries using spatial relationships between features.

Table Joins

- Properties → Joins
- Processing Toolbox → Vector General → Join Attributes by Field

69

This slide describes how to perform a table join in QGIS.

Spatial Join

- Processing Toolbox → Vector General → Join Attributes by Nearest
- Processing Toolbox → Vector General → Join Attributes by Location

70

This slide lists tools for performing spatial joins in QGIS.

QGIS Queries 1

- Use the utah project.
- Use queries to answer the following questions.

1. How many counties in Utah do not have an airport or airfield?
2. How many points in the towns layer is of the City TYPE?
3. How many points in the towns layer have a population larger than 5,000 (POPULATION field)?
4. How many HUC12 watersheds occur within the Raft HUC8 watershed?
5. How many of the trail features are Designated as pedestrian?

71

These exercises ask you to answer questions using queries in QGIS.

QGIS Queries 2

- Use the utah project.
- Use queries to answer the following questions.

1. How many units in the geology layer have a MAJOR1 type of Limestone?
2. How many counties contain some limestone geology?
3. How many towns have a trail within 1 kilometer?

72

This is the end of this lecture module.

Please return to the West Virginia View Webpage for additional content.

Thanks! Hope you found this useful.